Statistical Inference for Scientific Instruments: Event Analysis for the Gamma-ray Large Area Space Telescope. Grant Number NNG05GC80G – Annual Report, Year 2

Robin D. Morris, USRA-RIACS
Johann Cohen-Tanugi, SLAC

Abstract

We report on progress on developing a new event analysis methodology for the Large Area Telescope (LAT) instrument on the forthcoming Gamma-ray Large Area Space Telescope (GLAST) mission. In year 1 we demonstrated the basic methodology on simulations of muons incident on the LAT, and began integrating the methodology into the GLEAM (GLAST Event Analysis Machine) software framework. Work in year 2 has progressed on to the development of the methodology to analyze electron events. These are more complex than muons because of the significant probability of the production of secondary photons. The new methodology has been shown to be able to infer the production of photons that are not explicitly detected anywhere in the LAT.

1 Introduction

The goal of this project is to develop a fully probabilistic reconstruction methodology for the Large Area Telescope (LAT) instrument on GLAST. The LAT is a pair-conversion instrument, where incident gamma-rays are converted into electron-positron pairs, and these charged particles are tracked as they traverse the active silicon microstrip layers in the tracker towers. We discussed in the report for year 1 how the statistics of the pair-production process require that, for an accurate determination of the incident direction of the photon, the energy of the electron and positron must be determined separately. This motivates the study of charged particles incident on the detector. In year 1 we studied the response of the LAT to incident muons. In year 2 we have focused on the response of the LAT to electrons.

When electrons are incident on the LAT, as they traverse the tungsten conversion layer they are subject to a number of physics processes that complicate the event reconstruction.

The major process is multiple scattering, by which the trajectories of the charged particles is changed, and by which they loose energy. They are also subject to Bremsstrahlung, whereby a charged particle, when accelerated in the charge field of an atom, has its direction changed, and also produces a secondary photon. The photon carries some of the energy of the charged particle, but will typically not be detected anywhere in the tracker. The energy carried by the photon can be considered to have "disappeared", and its presence can only be inferred by careful study of the behavior of the charged particle before and after the layer in which the Bremsstrahlung process occurred.

A number of other processes result in the production of secondary charged particles. Secondary charged particles significantly complicate the analysis, as they also cause the silicon microstrips to

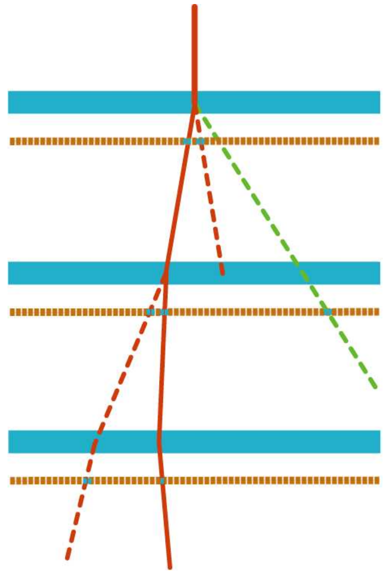

Figure 1: Schematic of an electron incident on the LAT. At the first layer a photon (green dashed line) and a secondary electron (red dashed line) are produced. At the second layer a further secondary electron is produced, and the first is absorbed in the tungsten foil.

fire, with the consequence that the trajectories of the initial charged particles and the secondary charged particles must be disambiguated.

Figure 1 shows a schematic of an electron incident on the LAT, and the resulting production of photons and secondary charged particles. Remember, however, that we do not observe the actual trajectories, rather, we only observe the hits on the silicon microstrips (indicated by the short blue lines in the orange layers below the blue tungsten layers). It is from these hits that we must infer the structure of the event, and there will necessarily be some ambiguity as to the precise structure of the event – by considering only the hits on the microstrips we can hypothesize a large number of events that are consistent with the pattern of microstrips that fired. The goal of this project is to generate these hypotheses, and compute the relative probability of each, to determine probabilistically the most likely structures of the actual event.

2 Electrons Incident on the LAT

When an electron is incident on a tungsten foil in the LAT, the major physics processes are

- Multiple Coulomb Scattering (MSC), which always occurs.
- Production of a secondary photon, which occurs approximately 25% of the time for a 100MeV electron.
- Production of a secondary photon, which occurs approximately 1% of the time for a 100MeV photon.

The probabilities of each secondary producing process are weak functions of energy in the range considered[1]. We concentrate here on the production of secondary photons, and neglect for now the production of secondary electrons.

Secondary photons carry away energy of the electron, and this energy is typically not detected anywhere in the LAT. The physics of the Bremsstrahlung process tells us, however, the distribution of energy in a secondary photon, given that one has been produced. This energy distribution is proportional to 1/e, where e is the energy of the incident electron. Because it is unknown whether a photon has been produced, the energy loss at any layer is made up of two components, 1) the energy loss due to multiple scattering, which has a Landau distribution, and 2) the energy loss due to a possible secondary photon. The total energy loss is a mixture, with the first component being Landau, and the second being the convolution of the Landau distribution with the 1/e distribution, weighted by the probabilities of not producing/producing a secondary photon. It must be noted that this distribution has a very long tail, up to the energy of the incident electron.

We implemented a `tool` within the GLEAM software framework to compute the energy loss distribution for an electron traversing the tungsten foils, based on statistics gathered from expensive simulations of electrons incident on the LAT foils. This tool, called `ElectronScatteringTool`, implements the interface defined in `IScatteringTool`. It computes the probability of a specified energy loss occurring, given the energy of the incident electron, and also computes the probability of a given scattering angle, again conditioned on the energy of the incident electron. The tool was integrated into the `CalSampler` framework that we are developing, and the parameters of the MCMC algorithm needed to work effectively with electrons were determined.

Results

Figure 2 shows the analysis of a 100MeV electron incident on the LAT. The top left panel shows the simulated event in the FRED display. The electron is incident at the top of the LAT, and produces a number of secondary photons as it traverses the instrument. The top right panel shows the estimated energy of the event. The energy estimate is peaked around 65MeV. While this is somewhat different to the electron's actual energy of 100MeV, it is not of great concern as the data available to estimate the electron's energy is very sparse – it is detected at most 12 times as it traverses the LAT, and estimating its energy, as has been previously discussed, can be compared with estimating the variance of a Gaussian distribution from samples from that distribution. With such a small number of samples, we expect that some of the estimates will not be close to the simulated value. This will be especially true when high-energy secondary photons are produced, as is the case here.

The bottom left panel shows the estimate of the energy lost by the electron at layer 9. This distribution is very close to the Landau distribution for the energy loss by multiple scattering, indicating that at this layer, multiple scattering was the dominant process. The bottom right panel shows the energy lost by the electron at layer 10. At this layer the estimate of the energy lost is significant, with a peak at around 35MeV. The event reconstruction is telling us that a secondary photon was produced at this layer, and that the energy of the secondary photon was likely to be around 35MeV. Inspection of the simulation of the event in the top left panel shows that a photon was indeed produced at that layer (it is circled in red on the figure). Looking in detail at the

[1]We are concerned here with particles in the energy range of 30MeV-500MeV. Above that range the multiple scattering is so small that the trajectory of the electron is to all intents and purposes a straight line, and its energy cannot be estimated in the tracker.

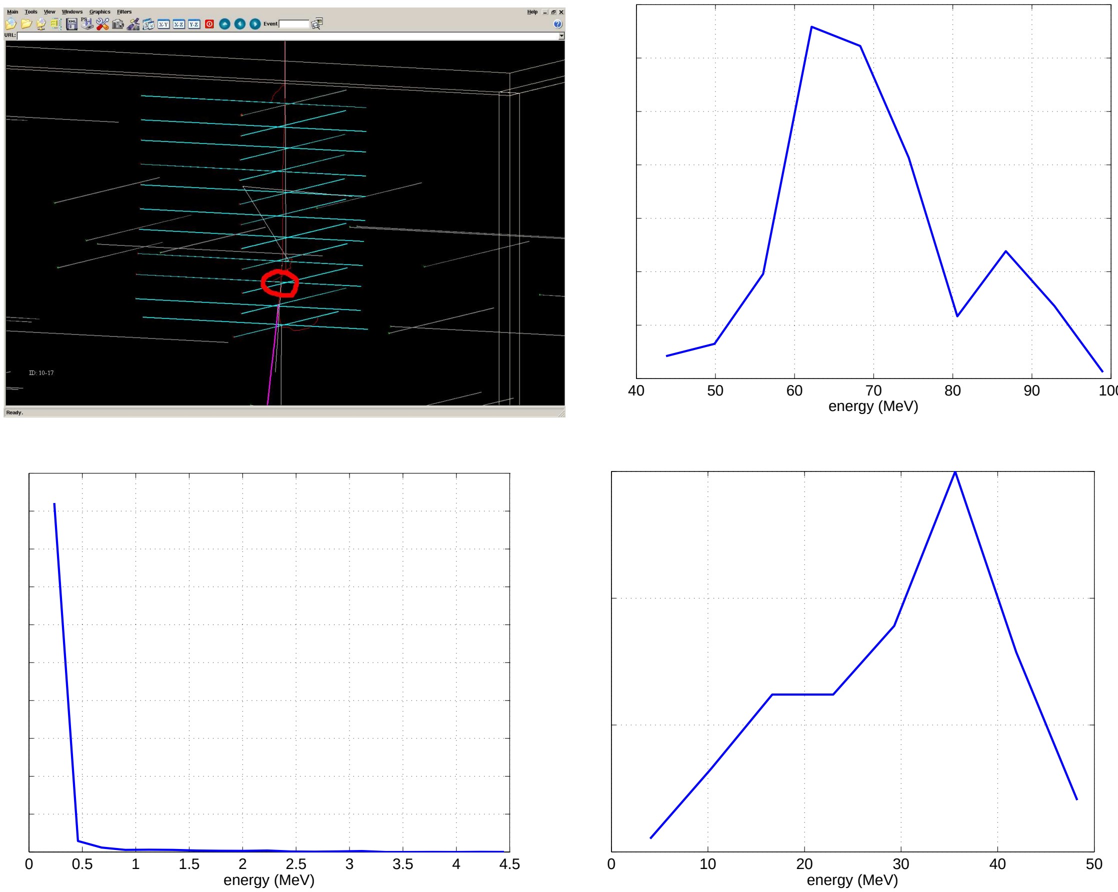

Figure 2: Results of the analysis of a 100MeV electron incident on the LAT. Top left - the simulated event; top right - the estimate of the energy of the electron; bottom left - the estimate of the energy lost by the electron in the tungsten foil at layer 9; bottom right - the estimate of the energy lost by the electron in the tungsten foil at layer 10. Note in the illustration of the simulated event that a photon was produced as indicated in the red circle.

simulation data for this event, the photon in fact had an energy of 80MeV. Note that the electron was detected only twice after the photon was produced, so there is extremely little data on which the estimate of the energy loss can be made. It is therefore surprising that the photon was detected at all, and not surprising that it's energy was not determined 100% accurately.

3 Work Plan for Year 3

We present the work plan for the final year of the project, which outlines the stages required to complete a system which will analyze gamma-ray events in their full generality. We will also complete extensive testing and documentation of the system, and write and submit papers to both physics and statistics journals describing the analysis methodology and results.

1. Determination of the Configurations for Incident Electrons (2 months)
   (a) Develop a tool to enumerate the possible configurations of secondary particles that could have resulted in the observed microstrip firings.
2. Determination of the Configurations' Probabilities (2 months)
   (a) Development of a tool that uses the output of the MCMC sampling and evaluates the relative probability of each of a set of configurations.
3. Comparative Testing (1 month)
   (a) Test the electron estimation algorithm on simulated electrons, and compare with the conventional estimation.
   (b) If possible, test on beam-test data from the actual instrument.

   Milestone - May 2007: A paper describing the methodology and results for the analysis of electrons incident on the LAT.
4. Determination of the Configurations for Incident Photons (2 months)
   (a) Expand the functionality of the previous tool to also enumerate the possible configurations for an incident photon.
5. Analysis of Specified Gamma-ray Configurations (2 months)
   (a) Expand the functionality of the previous tool to also include the two primary charged tracks and associated secondaries.
6. Determination of the Configurations' Probabilities (2 months)
   (a) Expand the functionality of the previous tool to also evaluate the relative probabilities for configurations corresponding to photons incident on the detector.

   Milestone - November 2007: Completion of the code for event analysis.
7. Complete the documentation (1 month)
   (a) Complete the documentation of the system.

(b) Prepare journal papers (for both statistics and physics journals) describing the analysis system.

8. Analysis of events from the actual instrument (1 month)

   (a) Attempt to obtain event data from the actual instrument.

   (b) Analyze this data using the new analysis framework. Resolve any problems that become apparent.

   (c) Demonstrate the framework to interested parties in the GLAST collaboration.

   (d) Write final report.

   Milestone - January 2008: Submission of journal papers and final report.

4 Other Related Activities

At the conference "Statistical Problem in Particle Physics, Astrophysics and Cosmology" that was attended by RDM in 2005, a paper was presented regarding the extraction of the CKM Phase α (a component of the Standard Model of particle physics) from the BaBar and Belle experiments. This paper put forward a fundamentalist frequentist view of the data analysis problem, and criticized another group that was applying Bayesian methodology to this problem. We are developing and applying Bayesian methodology to event analysis for GLAST, so this criticism was of great interest. A preprint of the paper was made available earlier this year (arXiv:hep-ph/0607246), and close study of this paper revealed a number of mis-understandings and mis-applications of the Bayesian methodology as applied to this problem in particle physics. We have written a technical report ("Bayesian Statistics That Works: How to Extract the CKM Phase α") showing how to correctly analyze this problem. We are currently awaiting feedback from a number of members of the particle physics community before submitting the paper for publication.

5 Dissemination of Results

5.1 Conference Presentations

- "Modern Statistical Methods for GLAST Event Analysis", invited talk at *Interface 2006*, 38th Symposium on the interface of statistics, computing science and its applications. May 24th-27th, 2006. Session organized by David van Dyk.

- "The Sub-atomic Particle Filter", presented at *Valencia 8* (8th Valencia International Meeting on Bayesian Statistics/World Meeting of the International Society for Bayesian Analysis), Benidorm, June 2006

- "Event Analysis for GLAST - a detailed statistical analysis", presented at *Statistical Challenges in Modern Astronomy IV*, Penn State, June 2006.

5.2 Conference Proceedings

- "Event Analysis for GLAST - a detailed statistical analysis", to appear in *Statistical Challenges in Modern Astronomy IV*, G.J. Babu and E. Feigelson (eds), Astronomical Society of the Pacific, 2007

5.3 Technical Reports

- "Bayesian Statistics That Works: How to Extract the CKM Phase α", Technical Report, RIACS, 2006

5.4 Seminars

- "Event Analysis for GLAST", Code S-Code T collaboration workshop, NASA Ames Research Center, March 2006
- "Two Statistical Problems in (astro-)Particle Physics", Department of Statistics, University of California, Santa Cruz, November 2006
- "Event Analysis for GLAST", Department of Engineering, Cambridge University, UK, December 2006

5.5 Conference Committees

- Member of the Technical Program Committee for EUSIPCO 2006 (14th European Signal Processing Conference.
- Member of the Technical Program Committee for SIGMAP 2006 and 2007